

10510677SEQ-72610  
SEQUENCE LISTING

<110> Sanofi Pasteur, Ltd.  
Therion Biologics, Inc.

<120> Modified CEA Nucleic Acid and Expression Vectors

<130> API-01-20-US

<140> 10/510,677

<141> 2003-04-09

<150> US 60/370,972

<151> 2002-04-09

<160> 33

<170> PatentIn version 3.5

<210> 1

<211> 3564

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (1663)..(1663)

<223> n is a, c, g, or t

<400> 1  
agcaggaccg gggcctgtgt cgctatgggt tcccccgccg ccccgaggg agcgctgggc 60  
tacgtcccgcg agttcactcg ccactccccc gacgtgtgg gcaacctcaa cgagctgcgc  
ctgcgcggga tcctcactga cgtcacgcgt ctgggtggcg ggcaacccct cagagcacac  
aaggcagttc tcatcgccgtc cagtggtttc tcattttccaa tttccgggg ccgtgcggga 120  
gtcgggggtgg acgtgtcttc tctgtccccc ggtcccgaaag cgagaggctt cggccctcta  
ttggacttca tgcgtacttc ggcgcgtgcgc ctctctccag ccactgcacc agcagtccata 180  
gcggccgcca cctatttgc gatggagcac gtgggtccagg catgccaccg ctccatccag  
gccagctatg aacctctggg catctccctg cggcccccgg aagcagaacc cccaaacacc 240  
ccaacggccc ctccaccagg tagtcccgagg cggtcccgaaag gacaccaggaa cccacact 200  
gaatctcgaa gctgcagtca aggccccccc agtccagcca gcccgtaccc caaggccctgc  
aactggaaaa agtacaagta catcggtcta aactctcgagg cctcccaagc agggagctgt 360  
gtcggggaga gaatttctgg tcaaccttgc ccccaagccca ggctcccgag tggagacgag  
gcctccagca gcagcagcagc cagcagcagc agcagcagtg aagaaggacc cattccgtgt  
ccccagagca ggtctctcc aactgtcgcc actgtcgagt tcaaatgtgg ggctcccgac 380  
agtacccctt acctccctac atcccgagct caagacaccc tggatcacc ctctgaacgg  
gctcgccac taccgggagt gaattttca gctgcaggaa ctgtggggct ttggccagggt 440  
900  
960

10510677SEQ-72610

gctcatcgaa	ggctggactc	cttgggtcct	ggggacgaa	acaaacccta	taagtgtcag	1020
ctgtgcccgt	cttcgttccg	ctacaaggc	aaccttgc	gtcaccgtac	agtgcacaca	1080
ggggaaaagc	cttaccactg	ctcaatctgc	ggagccgtt	ttaaccggcc	agcaaaccgt	1140
aaaacgcaca	gccgcaccca	ttcggggagag	aagccgtata	agtgtgagac	gtgcggctcg	1200
cgctttgtac	agggtggcaca	tctgcggg	cacgtgcgt	tccacaccgg	ggagaagccc	1260
tacccttgcc	ctacctgcgg	aaccgc	cgccac	agaccctaa	gagccacgtt	1320
cgcatccaca	ccggagagaa	gccttac	tgcacccct	gtgcgtca	tttccggcac	1380
aagagtcaac	tgcggctca	tctgcggc	aaacacggag	ctgctacca	caccaaagt	1440
cactaccaca	tttctggggg	gccttagt	agcgcaggcc	caggccccac	ttgcgttctg	1500
cgggtggaa	agtcgcaggc	ccaggc	cttccctatc	aggctggc	ataggggtgt	1560
gcaggccac	tttggtatca	gaaattgc	ccctttaat	tttctactg	ggagagcagg	1620
ggtggcagat	cctggctaga	tctgcct	ttttgtgtt	canaccctt	tccccacaag	1680
ccagattgtt	tctgaggaga	gagctagta	ggggctggg	aaggggagag	attggagttc	1740
tgtgtccct	aagggatag	ccctccac	gtggccccca	ttgcattc	tttatctgt	1800
aaataataatt	tattgaggcc	tttgggtggc	accgggg	tcattcgatt	gcatttccca	1860
ctcccttctt	ccacaagtgt	gattaaaatg	gaccaga	acagaagg	agatcacagc	1920
tctgtggca	gagattacta	gccttg	ctctgtt	gttgggtat	tttatattat	1980
ttctgtcata	acttttatct	tttagattgt	tcttct	gttgggttgc	ttgttagttt	2040
gtttaaaatg	aaaaaagggg	ttctctgt	tctgc	taattctagg	tctggAACCT	2100
ttattgttc	tagggcagct	ctggaa	catgc	ggggattgt	ggaaattgggt	2160
ctctgttatt	ctggatgtt	taggttct	agcag	tct	aaatggata	2220
ctgttctca	agggtgatag	gaaccattat	gttgagcc	aaatgg	aataataat	2280
gccttctgga	ggctgtgggt	gtgggg	ctgtatct	atccgtatc	actccaa	2340
gaggctgtgg	gtgtgggg	ttctgtatct	ggatcc	taactccaa	tggaggctgg	2400
cagg	ttttc	tgcaagatgg	tccagaatct	aaatgtccc	attatctgg	2460
ttggctctgc	tgtatccatc	tatagtggta	gagacc	accaggctcaag	tggagtccat	2520
catcccttca	cgggggc	ttcttagtac	tgagtgt	atccgtatgg	ggagagatca	2580
gacattcctt	attcagagat	atgtgac	ttctgt	actc	tatgaatgtt	2640
atggccttag	gaagaatcat	gaaactctt	agctt	gatgtaa	atgtttaacc	2700
catcccttac	tacagaggca	tatgggtt	aatgttac	gggg	tcttctctt	2760
gagcccttc	ttcc	tttagt	gggtttgg	catcttct	ctattgagtt	2820
cccttcttc	ctctagaagg	gtgg	gtt	ggtaacc	ttttaaaa	2880

10510677SEQ-72610

tgcacccatggtc	ttccccatccc	tgccatccctg	tctggaaacca	gtaatgcacat	tagaacccctc	2940
cataggaaaa	gaaaaggggc	tgaggctccat	tctgggtttt	ctgttagttt	gttgggatttt	3000
tttgtggcat	tacagatgtt	aaagattgac	tagcccatag	gccaaaggcc	tgttctagtt	3060
gaccaagttt	caagtaggat	taagagggtt	gttgggggt	gcagtttctg	gtgttaggcac	3120
ggtaggtaga	aagtggagaa	cagggttgc	tcttggctgg	gttggagtctc	tgaaatgtta	3180
gaagaagcgc	tgaaggctt	attgtatgtt	ctggccctt	ttggccctggg	gctttatctga	3240
ttatgggacg	agggttagaa	gtaagaagca	cttttggattt	tgtggggtag	aacttcaaca	3300
ataagtctgt	tcttagtggct	gtcgccctggg	gactagttag	aaagctactc	tttccctct	3360
tccctctttc	tccccatggc	cccaactgcag	aattaaagaa	ggaagaaggg	aaggcggagg	3420
agtctataag	aaggaatcat	gatttctatt	tagcagattt	gatggggcagg	tggagaatgc	3480
ctgggggttag	aaatgttaga	tcttgcacaca	tcagatcctt	ggaataaaaga	agccctctcg	3540
cocaaaaaaaaa	aaaaaaaaaa	aaaa				3564

<210> 2  
<211> 1440  
<212> DNA  
<213> *Homo sapiens*

10510677SEQ-72610

cctggggacg aagacaaacc	ctataagtgt	cagctgtgcc	ggcttcgtt	ccgctacaag	1020	
ggcaacctg	ccagtcatcg	tacagtgcac	acagggaaa	agccttacca	ctgctcaatc	1080
tgcggagccc	gtttaaccc	gcacgcaac	ctgaaaacgc	acagccgat	ccattcggya	1140
gagaagccgt	ataagtgtga	gacgtgcggc	tcgcgcctt	tacagggtgc	acatctgcgg	1200
gcgcacgtgc	tgtatccacac	cgggagaag	cccttacccctt	gccttacccgt	cgaaacccgc	1260
ttccgcacc	tgcagaccct	caagagccac	gttcgcacatcc	acaccggaga	gaagccttac	1320
cactgcgacc	cctgtggcct	gcattccgg	cacaagagtc	aactgcggct	gcatctgcgc	1380
cagaacacg	gagctgctac	caacaccaa	gtgcactacc	acattctcg	ggggccctag	1440

<210> 3  
<211> 65  
<212> DNA  
<213> Homo sapiens

<400> 3	atacccggaa	ctccctaagc	cttctattag	ctccaataat	agtaagcctg	tcgaagacaa	60
agatg							65

<210> 4  
<211> 70  
<212> DNA  
<213> Homo sapiens

<400> 4	gcctgtgtcc	ccttagactcc	aactcagcaa	cggaaataga	actctgaccc	tgtttaacgt	60
gaccaggAAC							70

<210> 5  
<211> 70  
<212> DNA  
<213> Homo sapiens

<400> 5	acgtgtttta	cggacccgat	gctcctacaa	tcagccctct	aaacacaagc	tatagatcag	60
ggggaaatct							70

<210> 6  
<211> 70  
<212> DNA  
<213> Homo sapiens

<400> 6	acgttaaacaa	gggtcagagt	tctatttccg	ttgctgagtt	ggagtctagg	ggacacaggc	60
agggactgtt							70

<210> 7

10510677SEQ-72610

<211> 70  
<212> DNA  
<213> Homo sapiens

<400> 7  
ctgatctata gcttgtgtt agagggctga ttgttaggac atcgggtccg taaagcacgt 60  
tgagaatcac 70

<210> 8  
<211> 63  
<212> DNA  
<213> Homo sapiens

<400> 8  
gatccactat tggcacggt aatattggga atgaacagtt cctgggtgga ctgttggaaa 60  
gtg 63

<210> 9  
<211> 70  
<212> DNA  
<213> Homo sapiens

<400> 9  
gacacagcaa gctacaaaatg cgaaacccaa aatccagtca gcgcaggag gtctgattca 60  
gtgattctca 70

<210> 10  
<211> 70  
<212> DNA  
<213> Homo sapiens

<400> 10  
tgaatcagac ctcctggcgc tgactggatt ttgggtttcg cattttagc ttgctgtgtc 60  
gttcctggtc 70

<210> 11  
<211> 79  
<212> DNA  
<213> Homo sapiens

<400> 11  
gatcctacac gtgccaagct cacaatagcg acacccgact caaccgcaca accgtacga 60  
cgattaccgt gtatgccga 79

<210> 12  
<211> 70  
<212> DNA  
<213> Homo sapiens

<400> 12  
catcctcaac tgggttagaa ttgttactag ttatgaatgg ttttggtggc tcggcataaca 60  
cggttaatcgt 70

<210> 13	
<211> 80	
<212> DNA	
<213> Homo sapiens	
<400> 13	
ttcttaaccca gttgaggatg aggacgcagt tgcattaact tgtgagccag agattcaaaa	60
taccacttat ttatggtggg	80
<210> 14	
<211> 80	
<212> DNA	
<213> Homo sapiens	
<400> 14	
gtctaatgtt aaccgcacat tgacactcct gtccgttact cgcaatgtat taggaccta	60
tgagtgtggc attcagaatgt	80
<210> 15	
<211> 80	
<212> DNA	
<213> Homo sapiens	
<400> 15	
tttgttatggc ccagacgacc caactatatac tccatcatac acctactacc gtccggcgat	60
gaacttgagc ctttcttgcc	80
<210> 16	
<211> 80	
<212> DNA	
<213> Homo sapiens	
<400> 16	
tgtatggaaac attcagcagc atactcaaga gttatttata agcaacataa ctgagaagaa	60
cagcggactc tatacttgcc	80
<210> 17	
<211> 80	
<212> DNA	
<213> Homo sapiens	
<400> 17	
taaaaacaata actgtttccg cggagctgcc caagccctcc atctccagca acaactccaa	60
acccgtggag gacaaggatg	80
<210> 18	
<211> 80	
<212> DNA	
<213> Homo sapiens	
<400> 18	

10510677SEQ-72610  
atgtgcggtt atcattagac aactgcaagc gtgggctaac cggcaaactt tggttattga 60  
cccaccataa ataagtggta 80

<210> 19  
<211> 80  
<212> DNA  
<213> Homo sapiens

<400> 19  
ggtcgtcttg gccataaaaa acattaagga taacagggtc ggagtgtatca acggataatt 60  
cattctgaat gccacactca 80

<210> 20  
<211> 80  
<212> DNA  
<213> Homo sapiens

<400> 20  
gctgctgaat gttccatca atcagccagg agtactgtgc aggggggttg gatgctgcat 60  
ggcaagaaag gctcaagtcc 80

<210> 21  
<211> 80  
<212> DNA  
<213> Homo sapiens

<400> 21  
cgaaacagt tattgttta actgttagtcc tgctgtgacc actggctgag ttattggctt 60  
ggcaagtata gagtccgctg 80

<210> 22  
<211> 47  
<212> DNA  
<213> Homo sapiens

<400> 22  
cctcagggttc acaggtaag gccacagcat ccttgtcctc cacgggt 47

<210> 23  
<211> 2106  
<212> DNA  
<213> Homo sapiens

<400> 23  
atggagtctc cctcgcccc tccccacaga tggtgcatcc cctggcagag gctctgtctc 60  
acaggctcac ttcttaacccctt ctggAACCCG cccaccactg ccaagctcac tattgaatcc 120  
acggccgttca atgtgcgaga ggggaaggag gtgcttctac ttgtccacaa tctggcccccag 180  
catcttttgc gctacagctg gtacaaagggt gaaagagtgg atggcaaccg tcaaattata 240  
qqatatqtaa taqqaactca acaaqctacc ccaqqqcccq catacagtqq tcqqaqata 300

10510677SEQ-72610	
atataccccca	360
atgcatccct	
gctgatccag	
aacatcatcc	
agaatgacac	
aggattctac	
accctacacg	420
tcataaagtc	
agatcttgc	
aatgaagaag	
caactggca	
gttccggta	
tacccggcgc	480
tgcccaagcc	
ctccatctcc	
agacaacaact	
ccaaaaccgt	
ggaggacaag	
gatgtgtgg	540
ccttcacctg	
tgaacctgag	
actcaggacg	
caacctacct	
gtgggtggta	
aacaatcaga	600
gcctcccgtt	
cagtcacccgg	
ctgcagctgt	
ccaaatggca	
caggaccctc	
actctattca	660
atgtcacaag	
aaatgacaca	
gcaagctaca	
aatgtgaaac	
ccagaacccca	
gtgagtgc	720
ggcgcagtg	
ttcagtcac	
ctgaatgtcc	
tctatggcc	
ggatgcccc	
accatttccc	780
ctctaaacac	
atcttacaga	
tcagggaaa	
atctgaacct	
ctccctgccac	
gcagcctcta	840
accacacctgc	
acagtaactct	
tggttgtca	
atgggacttt	
ccagcaatcc	
acccaagacg	900
tctttatccc	
caacatca	
gtgataata	
gttggatcta	
tacgtccaa	
gcccataact	
cagacactgg	
cctcaatagg	
accacagtca	
cgacgatcac	
agtctatgag	
ccacccaaac	960
ccittcatcac	
cagcaacaac	
tccaaaccccg	
tggaggatga	
ggatgctgt	
gccttaacct	1020
gtgaaacctg	
gattcagaac	
acaacctacc	
tgtgtgtgg	
aaataatcag	
agcctcccg	1080
tcaagtcccg	
gctgcagctg	
tccaaatgaca	
acaggaccct	
cactctactc	
agtgtcaca	1140
ggaatgatgt	
aggaccctat	
gagtgtggaa	
tccagaacga	
attaatgttt	
gaccacagcg	1200
acccagtc	
cctgaatgtc	
ctctatggcc	
cagacgaccc	
caccatttcc	1260
ccctcataca	
cttattaccg	
tccaggggtg	
aaacctcage	
tctctgtcc	
tgccgcctt	
aacccacctg	1320
cacagtattc	
ttggctgatt	
gatggaaaca	
tccagcaaca	
cacacaagag	
ctctttatct	1380
ccaaacatcac	
tgagaagaac	
agcggactct	
atactgtcc	
ggccaataaac	
tcagccatgt	1440
gccacagcg	
gactacagtc	
aagacaatca	
cagtctctgc	
ggagctgccc	
aagccctcca	1500
tctccagcaa	
caactccaaa	
cccggtggagg	
acaaggatgc	
tgtggccttc	
acctgtgaa	1560
ctgaggctca	
gaacacaacc	
tacctgtgg	
gggtttatgg	
tcaagcctc	
ccaggctgca	1620
gctgtccaa	
ggcaacagga	
ccctcactct	
attcaatgtc	
acaagaaatg	1680
acgcaagacg	
ctatgtatgt	
ggaatccaga	
actcagtgag	
tgcacaaaccgc	
agtgtacccag	1740
tcacccctgg	
tgtcccttat	
ggggccggaca	
ccccccatcat	
ttccccccca	
gactctgttt	1800
acccctgggg	
agcggaccc	
aacctctct	
gcccactcg	
ctcttaacccca	
tcccccggat	1860
attcttggcg	
tatcaatggg	
ataccgcgc	
aacacacaca	
agttctcttt	
atcgccaaa	1920
tcacgc	
taataacggg	
acctatgcct	
gttttgtctc	
taacttgg	
actggccca	1980
ataattccat	
agtcaagacg	
atcacagtct	
ctgcacatctgg	
aacttctct	
gggtctctcg	2040
ctggggccac	
tgtcgccatc	
atgattggag	
tgtgtgtttgg	
ggttgtctcg	
atata	2100
atata	2106

10510677SEQ-72610

<210> 24  
<211> 47  
<212> DNA  
<213> Artificial sequence  
<220>  
<223> Homo sapiens  
  
<400> 24  
ggacggtagt aggtgtatga tggagatata gttgggtcgt ctgggcc 47  
  
<210> 25  
<211> 27  
<212> DNA  
<213> Artificial sequence  
<220>  
<223> Homo sapiens  
  
<400> 25  
cagaatgaat tatccgttga tcactcc 27  
  
<210> 26  
<211> 45  
<212> DNA  
<213> Artificial sequence  
<220>  
<223> Homo sapiens  
  
<400> 26  
cgtgacgacg attaccgtgt atgagccacc aaaaccattc ataac 45  
  
<210> 27  
<211> 45  
<212> DNA  
<213> Artificial sequence  
<220>  
<223> Homo sapiens  
  
<400> 27  
gttatgaatg gtttggtagt ctcatacacg gtaatcgtag tcacg 45  
  
<210> 28  
<211> 2106  
<212> DNA  
<213> Artificial sequence  
<220>  
<223> Homo sapiens  
  
<400> 28  
atggagtc tcccgcccccc tccccacaga tggtgcatcc cctggcagag gctcctgctc 60  
acagcctcac ttcttaacctt ctggAACCCG CCCACCACTG CCAAGCTCAC TATTGAATCC 120  
acGCCGTTCA ATGTCGAGA GGGGAAGGG AGTGTCTAC TTGTCACAA TCTGCCAG 180  
CATCTTTTG GCTACAGCTG GTACAAAGGT GAAAGAGTGG ATGGCAACCG TCAAATTATA 240  
GGATATGTAAG TAGGAACTCA ACAAGCTACC CCAGGGCCCG CATACTGG TCGAGAGATA 300  
ATATACCCCA ATGCATCCCT GCTGATCCAG AACATCATCC AGAATGACAC AGGATTCTAC 360

## 10510677SEQ-72610

accctacacg tcataaaagtc agatcttgtg aatgaagaag caactggcca gttccgggta	420
tacccggaaac tccctaagcc ttcttattagc tccaaataata gtaaggctgt cgaagacaaa	480
gatggcgctcg cttttacatg cgagccccaa actcaagacg caacatattt ctggtgggtg	540
aacaaccagt ccctgcctgt gtcctttaga ctccaaactca gcaacggaaa tagaactctg	600
accctgttta acgtgaccag gaacgacaca gcaagctaca aatgcgaaac cccaaatcca	660
gtcagcgcca ggagggtctga ttcatgtattt ctcacagtgc ttacggacc cgatgctcct	720
acaatcagcc ctctaaacac aagctataga tcagggaaa atctgaatct gagctgtcat	780
gccgcctagca atccctccgc ccaatacagc tggtttgtca atggcacttt ccaacagtcc	840
accaggaaac tgttcattcc caatattacc gtacaataa gtggatccca cacgtgccaa	900
gctcacaataa ggcacaccgg actcaaccgc acaaccgtga cgacgattac cgtgtatgag	960
ccacccaaac cattcataac tagtaacaat tctaaccag ttgaggatga ggacgcagtt	1020
gcattaaactt gtgagccaga gattaaaaat accacttattt tatgggtgggta caataaccaa	1080
agtttgcccg tttagccacg ctgcagtttgc tctaattgtata accgcacatt gacactctg	1140
tccgttactc gcaatgatgtt aggcacccat gagtgtggca ttcaaatgttattatccgtt	1200
gatcactccg accctgttat ccttaatgtt ttgtatggcc cagacgaccc aactatatact	1260
ccatcataaca cctactaccg tcccgccgtg aacttgcgca ttcttgccca tgccatcc	1320
aaccccccgtg cacagttactc ctggctgtt gatggaaaca ttcaatgtata accgcacattt	1380
ttattttataa gcaacataac tgagaagaac agcggactct atacttgcata ggcataaaac	1440
tcagccactg gtccacagcag gactacagtt aaaaacaataa ctgtttccgc ggagctgccc	1500
aaggccctca tctccagca caactccaaa cccgtggagg acaaggatgc tggcccttc	1560
acctgtgaac ctggggctca gaacacaacc tacctgtgtt gggtaatgg tcagagccctc	1620
ccaggctgcgactc ccaggctgcgactc ccgttccaaat ggcacacagg ccctactctt attcaatgtc	1680
acaagaaatg acgcaagacg ctatgtatgtt ggaatccaga actcgtgatgc tgccaaaccgc	1740
agtgaccccg tcaaccctggatgttccat gggccggaca cccccatcat ttccccccca	1800
gactcgtt acctttcggtt agcggacccctt aacctcttccctt gccactcggtt ctctaaacccca	1860
tccccccgtt attcttggcg tataatggg ataccgcgc aacacacaca agttctctttt	1920
atcgccaaa tcaacgcggaaataa acctatgcctt gttttgtctc taacttggctt	1980
actggccgcataaattccat agtcaagacg atcacagtctt ctgcgtctgg aacttctctt	2040
ggctctcttagt ctggggccac tgcggccatc atgatggag tgctgggttgg gggttgcgtcg	2100
atata	2106

10510677SEQ-72610

<211> 35  
<212> DNA  
<213> Artificial sequence  
<220>  
<223> Homo sapiens

<400> 29  
ctggcgcc ttctttattc tatacttaaa aagtg 35

<210> 30  
<211> 36  
<212> DNA  
<213> Artificial sequence  
<220>  
<223> Homo sapiens

<400> 30  
ctgggttaccag aaaaactata tcagagcaac cccaac 36

<210> 31  
<211> 26  
<212> DNA  
<213> Artificial sequence  
<220>  
<223> Homo sapiens

<400> 31  
ttggatccat ggagtctccc tcggcc 26

<210> 32  
<211> 26  
<212> DNA  
<213> Artificial sequence  
<220>  
<223> Homo sapiens

<400> 32  
ttggatccct atatcagagc aacccc 26

<210> 33  
<211> 16  
<212> PRT  
<213> Homo sapiens

<400> 33

Ser Arg Arg His His Cys Arg Ser Lys Ala Lys Arg Ser Arg His His  
1 5 10 15